CHEST

Original Research

SIGNS AND SYMPTOMS OF CHEST DISEASES

Dyspnea-12 Is a Valid and Reliable Measure of Breathlessness in Patients With Interstitial Lung Disease

Janelle Yorke, PhD; Jeffrey Swigris, DO, MS; Anne-Marie Russell, MSc; Shakeeb H. Moosavi, PhD; Georges Ng Man Kwong, MD; Mark Longshaw, MD; and Paul W. Jones, PhD

Objective: **In this study, we aimed to determine the validity and reliability of the Dyspnea-12 questionnaire (D-12) for the assessment of breathlessness in patients with interstitial lung disease (ILD).**
Methods: **A total of 101 patients with ILD completed the D-12 (scale range, 0-36, with a high score indicating worse dyspnea), Medical Research Council (MRC) dyspnea scale, St. George Respiratory Questionnaire (SGRQ), and Hospital Anxiety and Depression Scale (HADS) at baseline, and 84 patients completed the D-12 and a global health transition score at follow-up 2 weeks later. D-12 psychometric properties, including floor and ceiling effects, internal consistency, test-retest reliability, and construct validity were examined.**
Results: **The D-12 showed good internal consistency (Cronbach α, 0.93) and repeatability (intraclass correlation coefficient, 0.94). Its scores were significantly associated with MRC grade ($r = 0.59$; $P < .001$), SGRQ (symptoms, $r = 0.57$; activities, $r = 0.78$; impacts, $r = 0.75$; total, $r = 0.79$; $P < .001$). Factor analysis confirmed the previously determined structure of the D-12 in this patient group.**
Conclusion: **In patients with ILD, the D-12, a patient-reported measure of dyspnea severity that requires no reference to activity, is a reliable and valid instrument. It is short, simple to complete, and easy to score.** ***CHEST 2011; 139(1):159–164***

Abbreviations: 6MWD = 6-min walk distance; CFA = confirmatory factor analysis; CFI = comparative fit index; D-12 = Dyspnea-12 questionnaire; DLCO = diffusing capacity of the lung for carbon monoxide; HADS = Hospital Anxiety and Depression Scale; HRQL = health-related quality of life; ILD = interstitial lung disease; IPF = idiopathic pulmonary fibrosis; MRC = Medical Research Council; NFI = nonfit index; NNI = nonnormed index; SGRQ = St. George Respiratory Questionnaire

Interstitial lung disease (ILD) refers to a cluster of fibroinflammatory conditions for which dyspnea is a cardinal symptom. For patients with purely fibrotic ILD, currently available medical treatments have had no impact on survival.[1] In patients with ILD, measures of health status and symptom perception are becoming important methods for assessing the impact of disease and treatment efficacy.[2,3] However, little progress has been made in the development of patient-reported outcomes designed for this patient group. Despite its prevalence and prominence, dyspnea associated with ILD has received little attention in clinical trials and studies exploring disease trajectory. This gap probably reflects the limited availability of

Manuscript received March 15, 2010; revision accepted June 8, 2010.
Affiliations: From the Faculty of Health and Social Care (Dr Yorke), School of Nursing, University of Salford, Greater Manchester, England; Interstitial Lung Disease Program (Dr Swigris), National Jewish Health, Denver, CO; University of Huddersfield (Ms Russell), West Yorkshire, England; Faculty of Medicine (Dr Moosavi), Imperial College London, London, England; Pennine Acute Trust (Drs Ng Man Kwong and Longshaw), Crumpsall, England; and St. Georges University of London (Dr Jones), London, England.
Funding/Support: This study was funded by Action Medical Research, England. Dr Swigris is supported in part by a Career Development Award from the National Institutes of Health (K23 HL092227).

Correspondence to: Janelle Yorke, PhD, School of Nursing & Midwifery, Faculty of Health and Social Care, Mary Seacole Building, Room 1.43, University of Salford, Greater Manchester, M66PU, England; e-mail: j.yorke@salford.ac.uk

DOI: 10.1378/chest.10-0693

information relating to dyspnea perception and the absence of robust data that point to the most valid and reliable criteria for assessing dyspnea in this patient group.

Several instruments are available to capture the effect of dyspnea on patients, although these largely have been developed and validated for use in patients with COPD and have been associated with activity limitation.[4,5] The Dyspnea-12 questionnaire (D-12) was developed using descriptors of breathlessness relevant to patients with a variety of cardiopulmonary diseases, including ILD.[6] It measures the current level of a patient's breathlessness severity, incorporating both physical and affective aspects, and does not depend on activity limitation. It has demonstrated validity in patients with COPD,[6] but it has not been tested for validity in patients with ILD. In this study, we aimed to determine the reliability and validity of the D-12 in patients with ILD.

Materials and Methods

Subjects

All patients with documented ILD attending specialist outpatient clinics in northwest England between February 2008 and August 2009 were asked to participate. A total of 120 eligible patients were identified, and 101 participated after providing informed written consent. The study was approved by the Salford and Trafford (Greater Manchester, North West, England) Research Ethics Committee (07/H1004/168).

Study Design

For test-retest reliability, participants completed a set of four questionnaires, including the D-12, in the clinic at baseline and at home 2 weeks later. This time period was considered to be long enough for participants not to recall (and simply reiterate) their baseline D-12 responses while their clinical condition remained constant. Pulmonary function tests and the 6-min walk distance (6MWD) test were conducted at baseline according to American Thoracic Society/European Respiratory Society guidelines[7,8] and percent-predicted FVC and diffusing capacity of the lung for carbon monoxide (DLCO) were expressed for age, sex, and height.

Questionnaires Completed

D-12: Participants completed the D-12 in reference to their experience of breathlessness "these days" at baseline and follow-up.[6] D-12 consists of 12 descriptor items on a scale of none (0), mild (1), moderate (2), or severe (3). It provides an overall score for breathlessness severity that incorporates seven physical items and five affective items. The time reference period for "these days" captures the current level of breathlessness experienced by patients as opposed to specifically on the day of the test or in response to a specific activity. Total scores from the D-12 range from 0 to 36, with higher scores corresponding to greater severity.

Hospital Anxiety and Depression Scale: Completed at baseline only, the 14-item Hospital Anxiety and Depression Scale (HADS) is a validated and widely used tool for assessing psychologic distress. The HADS comprises seven items that tap anxiety (score range, 0-21) and seven items that tap depression (score range, 0-21), with higher scores corresponding to greater distress.[9] For certain statistical tests, we considered a HADS anxiety score ≥ 10 to represent significant anxiety and a HADS depression score ≥ 10 to represent significant depression.[10] Because the D-12 contains items that reflect the psychologic consequences of breathlessness, the HADS was used to test any association between overall breathlessness severity and general psychologic well-being.

Medical Research Council Dyspnea Scale: Completed at baseline only, the Medical Research Council (MRC) dyspnea scale[4] (score range, 1-5, with higher scores indicating greater impairment) was used to classify participants according to activity limitation and to examine associations with D-12 scores.

St. George Respiratory Questionnaire: Completed at baseline only, the St. George Respiratory Questionnaire (SGRQ) is a widely used, standardized assessment[11] with three component scores: symptoms, activity, and impacts. Scores for each component and a total score calculated from the three components range from 0 to 100, with higher scores indicating greater disability. Although designed for patients with obstructive disease, the SGRQ has been found to be a valid measure of health-related quality of life (HRQL) in patients with restrictive disease.[3,12,13]

Global Health Transition Item: Completed at follow-up only, participants were asked whether their general health was much better, somewhat better, about the same, somewhat worse, or much worse than when they completed the questionnaires at baseline. The transition item was used to determine whether perceived general health condition had changed between the two time points.

Statistical Analysis

Psychometric properties, including floor and ceiling effects, internal consistency and test-retest reliability, and construct validity were examined. Statistical significance was set at $P \leq .05$. Parametric and distribution-independent tests were performed throughout the analyses. No major discrepancies between these two approaches were found, so results are reported from the parametric tests to allow for easier interpretation. Statistical tests were performed with SPSS, version 6 (SPSS Inc; Chicago, Illinois).

Reliability

Internal consistency was tested using Cronbach α for which values from 0.7 to 0.9 are believed to represent acceptable internal consistency for a multi-item instrument.[14] Test-retest reliability was assessed with the intraclass correlation coefficient in participants who recorded a global health transition of about the same.

Validity

Correlations between D-12 scores and SGRQ, HADS, MRC dyspnea grade, FVC%, DLCO%, and 6MWD were evaluated with Pearson coefficient. Associations with MRC dyspnea grade were calculated using one-way analysis of variance. Discriminative validity between different subgroups (eg, patients using supplemental oxygen vs those not using it) was tested using *t* test for unequal variance. Although D-12 provides a global score for dyspnea severity, we aimed to test whether the underlying patterning of items into physical and affective components that was reported previously[6] also was seen in this patient group. We used

the SAS PROC CALIS (SAS Institute, Cary, North Carolina) procedure to run confirmatory factor analysis (CFA). Five tests were applied to assess acceptable fit: χ^2 test, root mean square error of approximation, Bentler comparative fit index (CFI), Bentler and Bonnett nonnormed index (NNI), and Bentler and Bonnett nonfit index (NFI). Fit is indicated by a χ^2 $P \geq .05$, root mean square error of approximation < 0.06, CFI ≥ 0.9, NNI ≥ 0.9, and NFI ≥ 0.9.[15] See e-Appendix for details regarding CFA.

Results

Subject Details

A total of 101 subjects completed the questionnaires at study baseline and 84 at follow-up. Etiologies of ILD were idiopathic pulmonary fibrosis (IPF) (n = 67), asbestosis (n = 18), sarcoidosis (n = 10), connective tissue disease (n = 2), amiodarone toxicity (n = 2), pulmonary Langerhans cell histocytosis (n = 1), and nonspecific interstitial pneumonia (n = 1). The mean ± SD and range of the physiologic and patient-reported outcomes for the study population are summarized in Table 1.

D-12 Scores

There were no missing data for the D-12 at baseline or follow-up. The mean D-12 score for each time point was 16.1 ± 11.2 (range, 0-36) at baseline and 15.2 ± 10.9 (range, 0-36) at follow-up. There were no associations between the D-12 and age ($r = -0.08$; $P = .41$) or sex ($t = -0.42$; $P = .67$). The percentage of patients achieving the best score (floor effect) was 8%, and the percentage achieving the worst score (ceiling effect) was 2%.

Table 1—*Baseline Characteristics*

Characteristic	Value
Age, y	67 ± 10.9
Sex, %	
Female	70
Male	30
Race, %	
White	99
Black	1
FVC, % predicted	77 ± 19.5
FEV_1/FVC ratio	80 ± 10.9
Dlco, % predicted	51.6 ± 21
6MWD, m	337 ± 154
D-12 (baseline)	16.1 ± 11.1
MRC dyspnea grade 1-5	3.2 ± 1.3
SGRQ	
Symptoms	61 ± 23
Activity	65 ± 30
Impacts	41 ± 24
Total	53 ± 24
HADS	
Anxiety	8.9 ± 5.0
Depression	9.2 ± 5.4

Data are presented as mean ± SD, unless otherwise indicated. 6MWD = 6-min walk distance; D-12 = Dyspnea-12 questionnaire; Dlco = diffusing capacity of the lung for carbon monoxide; HADS = Hospital Anxiety and Depression Scale; MRC = Medical Research Council; SGRQ = St. George Respiratory Questionnaire.

D-12 Reliability

The internal consistency was high at baseline ($\alpha = 0.93$) and follow-up ($\alpha = 0.95$). The test-retest reliability was excellent at the 2-week follow-up (intraclass correlation, 0.94; $P < .001$) for patients whose global health score was unchanged.

D-12 Validity

Subjects taking oral steroid therapy at the time of the study (n = 32) had significantly higher D-12 scores compared with patients not taking steroids (n = 69) (mean D-12 score, 19.6 ± 8.2 vs 14.5 ± 11.4; $t = 2.3$; $P = .02$). Only 12 patients with IPF were receiving oxygen therapy at study entry. There was no difference in D-12 scores between subjects receiving oxygen therapy and those not (mean D-12 score, 16.1 ± 11.3 vs 15.8 ± 10.6; $t = 4.6$; P not significant).

Subjects with anxiety (HADS anxiety score, ≥ 10) reported more severe breathlessness than subjects who were not anxious (mean D-12 score, 19.2 ± 11.8 vs 13.4 ± 9.9; $t = 2.7$; $P = .009$). Subjects with depression (HADS depression score, ≥ 10) reported significantly more severe breathlessness than subjects who were not depressed (mean D-12 score, 19.3 ± 11.5 vs 13.5 ± 10.3; $t = 2.7$; $P = .008$). The D-12 correlated significantly with MRC dyspnea grade, HADS anxiety and depression scores, and SGRQ domains and total score (Table 2). Analysis of variance revealed a significant relationship with MRC dyspnea grade ($F = 16.1$; $P < .001$) (Fig 1). The D-12 correlated with Dlco% but not with FVC% (Table 3). Three CFA fit statistics, the CFI, NNI, and NFI, were all > 0.9, indicating acceptable fit and thus confirming the D-12 underlying structure in this patient group. See e-Appendix 1 for details regarding CFA.

Discussion

In this study, we examined the psychometric properties of the D-12 in a cohort of patients with ILD. We found the D-12 to possess excellent test-retest and internal consistency, and its scores correlated significantly and in the expected directions with other patient-reported outcome measures that assess HRQL or psychologic distress.

The results support the validity of the D-12 for assessing dyspnea at baseline and over time in this patient population. The D-12 is short and simple to use, attributes that most likely contributed greatly to

Table 2—*Associations Between the D-12 Other Self-Report Questionnaires at Study Entry (Pearson r)*

Questionnaire	Correlations With D-12 Scores
MRC dyspnea grade 1-5	0.59[a]
SGRQ	
Symptoms	0.57[a]
Activities	0.78[a]
Impacts	0.75[a]
Total	0.79[a]
HADS	
Anxiety	0.35[a]
Depression	0.22[b]

See Table 1 legend for expansion of abbreviations.
[a]$P < .001$.
[b]$P < .05$.

there being no missing data for it in our study. D-12 scores did not change among patients whose status has not changed, confirming its stability over time, and floor and ceiling effects were minimal for patients with mild-to-moderate ILD (the category to which physiologic measurements suggest our study cohort belonged).

CFA suggested that the previously identified underlying structure of the D-12 applies to this patient population. However, it should be noted that the D-12 was developed using Rasch methodology[16]; the 12 items were chosen because they conformed to the single construct being measured (ie, breathlessness severity), a requirement of the Rasch unidimensional model. As shown previously,[6] where items group appears to be based on severity associated with individual items as much as on any related underlying structure.

We found that D-12 scores were not affected by age or sex, which we expected because items with age or sex bias were removed in the D-12 development process. Rasch analysis was used in the development study to ensure that items possessed invariance across the severity range of breathlessness (calculated by patient responses to all items combined). Although there were no differences in pulmonary function between patients taking and those not taking oral corticosteroids, D-12 scores were significantly higher in patients taking them, suggesting that certain corticosteroid-related side effects contribute to higher D-12 scores, but this idea requires further testing. No difference was found in D-12 scores between patients receiving and those not receiving oxygen. In addition, post hoc analyses did not reveal any other differences in outcomes (including MRC scores and pulmonary function) between these groups. We believe that the small number of patients using supplemental oxygen probably contributed to a lack of power to detect a difference in D-12 scores between groups.

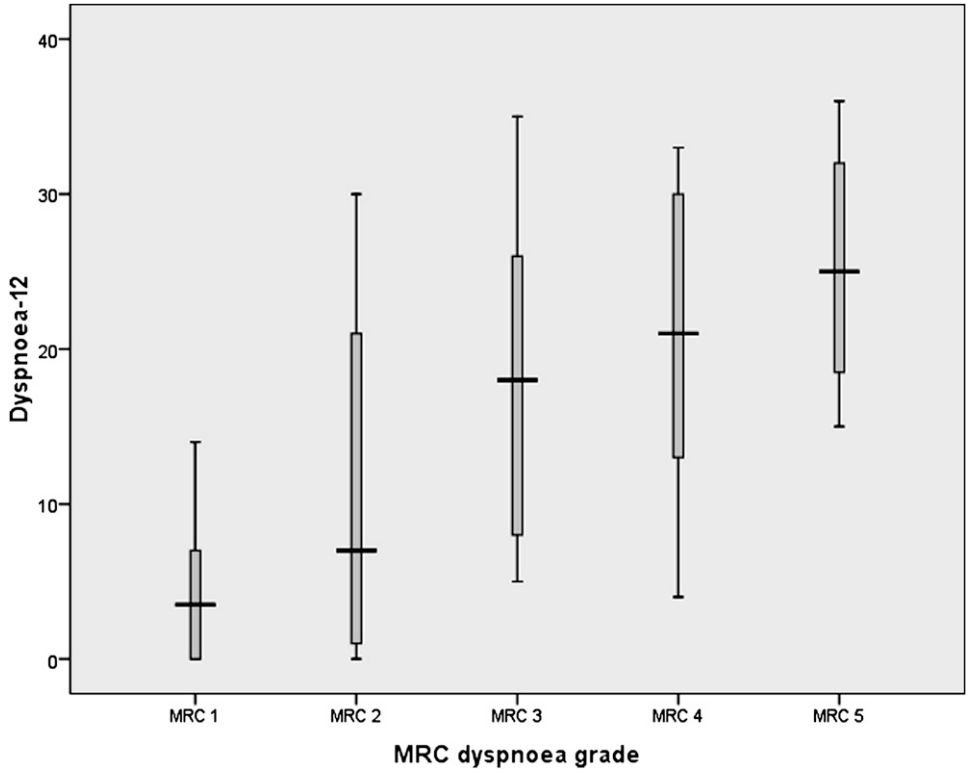

FIGURE 1. Dyspnea-12 questionnaire scores per MRC dyspnea grade. MRC = Medical Research Council.

Table 3—*Correlations Between Patient-Reported Outcomes and Physiologic Measurements (Pearson r)*

Physiologic Measurement	D-12	MRC	SGRQ Symptoms	SGRQ Activity	SGRQ Impacts
FVC%	−0.13	−0.21[a]	−0.13	−0.16	−0.24[a]
DLCO%	−0.27[a]	−0.36[b]	−0.16	−0.37[b]	−0.28[a]
6MWD, m	−0.51[b]	−0.59[b]	−0.32[b]	0.54[b]	−0.47[b]

See Table 1 legend for expansion of abbreviations
[a]$P < .05$.
[b]$P < .01$.

We observed significant correlations between D-12 scores and measures of activity limitation, psychologic distress, and HRQL scores. The strongest associations were with domains of the SGRQ. This finding is not surprising because the D-12 and SGRQ are both self-report measures, which have an inherently greater likelihood to be correlated than non-self-report measures, and dyspnea is the most important factor influencing HRQL in these patients.[12,13] There was a moderately strong relationship between D-12 score and MRC dyspnea grade, as represented by a progressively incremental increase in dyspnea severity per MRC grade. The correlation indicates that activity limitation caused by dyspnea (the construct that the MRC score measures) accounts for ~ 59% of the variability in D-12 scores. Thus, as it was designed to do, the D-12 captures much more about dyspnea than simply activity limitation. This notion is reinforced by the correlations between D-12 scores and anxiety and depression.

HADS scores of our patient cohort suggested anxiety and depression risk, taking into consideration that a score > 8 shows a risk.[9,12] This finding contrasts the study by Tzanakis et al,[12] who reported HADS anxiety and depression scores < 7 for a small cohort of patients with interstitial disease. Patients with higher anxiety and depression scores reported greater impairment on the D-12. Although significant, the strength of the relationships suggests that the D-12 and HADS measure different effects of well-being in patients with ILD. Specifically, the D-12 focuses on distress associated with breathlessness as opposed to a patient's general state of distress.

Functional capacity (as measured by the 6MWD test) seems to contribute more to patients' perceptions of dyspnea (as measured by the D-12) than FVC% or DLCO%; D-12 was not associated with FVC and weakly associated with DLCO. The D-12 is not a simple dyspnea index; it is not designed to simply rate degree of dyspnea for a given activity. In fact, it is not activity dependent. The D-12 measures general perceptions of breathlessness rather than in the context of specific activities and contains items that reflect the affective component of dyspnea, providing information beyond physiologic measurements such as FVC or DLCO.

This study has limitations. The majority of subjects had IPF, limiting the ability to generalize the results to patients with other types of ILD. The low number of patients who changed minimally limited our ability to reliably calculate D-12 responsiveness and minimum clinically important difference. We recommend that the D-12 be used in future interventional studies for ILD and that the minimum clinically important difference is calculated. However, despite these limitations and through rigorous statistical methodology, we observed the D-12 to be a reliable and valid measure of dyspnea in patients with ILD. We would encourage investigators to include the D-12 in future trials of novel therapies and other clinical studies of patients with ILD. It yields meaningful information that other dyspnea indexes do not.

In conclusion, patients with ILD report on the D-12—a measure of dyspnea severity that requires no reference to activity, unlike the majority of related instruments—is a reliable and valid instrument. It is short, simple to complete, and easy to score; therefore, we encourage investigators to include the D-12 as an outcome measure in future studies of patients with ILD.

Acknowledgments

Author contributions: *Dr Yorke:* contributed to the study concept and design, statistical analysis, and manuscript preparation. She had full access to the data and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Dr Swigris: contributed to the statistical analyses and manuscript preparation.
Ms Russell: contributed to the study concept and design, data collection, and manuscript preparation.
Dr Moosavi: contributed to the study concept and design.
Dr Ng Man Kwong: contributed to the data collection and manuscript preparation.
Dr Longshaw: contributed to the data collection and manuscript preparation.
Dr Jones: contributed to the study concept and design, statistical analyses, and manuscript preparation.
Financial/nonfinancial disclosures: The authors have reported to *CHEST* that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.
Other contributions: We thank all the patients and staff at Pennine Acute Trust who took time to take part in this study.
Additional information: The e-Appendix can be found in the Online Supplement at http://chestjournal.chestpubs.org/content/139/1/159/suppl/DC1.

References

1. Wells AU, Hirani N. The British Thoracic Society ILD Guidelines. *Thorax*. 2008;63(suppl 5):v1-v51.
2. Martinez TY, Pereira CAC, dos Santos ML, Ciconelli RM, Guimarães SM, Martinez JA. Evaluation of the short-form 36-item questionnaire to measure health-related quality of life in patients with idiopathic pulmonary fibrosis. *Chest*. 2000; 117(6):1627-1632.
3. Swigris JJ, Gould MK, Wilson SR. Health-related quality of life among patients with idiopathic pulmonary fibrosis. *Chest*. 2005;127(1):284-294.
4. Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax*. 1999; 54(7):581-586.
5. Mahler DA, Weinberg DH, Wells CK, Feinstein AR. The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest*. 1984;85(6):751-758.
6. Yorke J, Moosavi SH, Shuldham C, Jones PW. Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax*. 2010;65(1):21-26.
7. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med.* 2002;166(1):111-117.
8. American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). *Am J Crit Care Med*. 2000;161(2 pt 1): 646-664.
9. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand*. 1983;67(6):361-370.
10. Crawford JR, Henry JD, Crombie C, Taylor EP. Normative data for the HADS from a large non-clinical sample. *Br J Clin Psychol*. 2001;40(pt 4):429-434.
11. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis*. 1992;145(6):1321-1327.
12. Tzanakis N, Samiou M, Lambiri I, Antoniou K, Siafakas N, Bouros D. Evaluation of health-related quality-of-life and dyspnea scales in patients with idiopathic pulmonary fibrosis. Correlation with pulmonary function tests. *Eur J Intern Med*. 2005;16(2):105-112.
13. Nishiyama O, Taniguchi H, Kondoh Y, et al. Health-related quality of life in patients with idiopathic pulmonary fibrosis. What is the main contributing factor? *Respir Med*. 2005;99(4):408-414.
14. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297-334.
15. Suhr DD. Exploratory or confirmatory factor analysis? In: *Proceedings of the Thirty-first Annual SAS Users Group International Conference*. Cary, NC: SAS Institute; 2006: Paper 200-31.
16. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen, Denmark: Danish Institute of Educational Research; 1960.